P. Tung

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/211,691

DATE: 12/29/2000
TIME: 14:01:06

Input Set : A:\-129-1.app
Output Set: N:\CRF3\12292000\I211691.raw

```
 3 <110> APPLICANT: Gilbert, Michel
 4        Young, N. Martin
 5        Wakarchuk, Warren W.
 6        National Research Council of Canada
 8 <120> TITLE OF INVENTION: Fusion Proteins for Use in Enzymatic Synthesis of
 9        Oligosaccharides
11 <130> FILE REFERENCE: 019957-012910US
13 <140> CURRENT APPLICATION NUMBER: US 09/211,691
14 <141> CURRENT FILING DATE: 1998-12-14
16 <150> PRIOR APPLICATION NUMBER: US 60/069,443
17 <151> PRIOR FILING DATE: 1997-12-15
19 <160> NUMBER OF SEQ ID NOS: 18
21 <170> SOFTWARE: PatentIn Ver. 2.1
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 828
25 <212> TYPE: DNA
26 <213> ORGANISM: Neisseria meningitidis
28 <220> FEATURE:
29 <221> NAME/KEY: CDS
30 <222> LOCATION: (1)..(828)
31 <223> OTHER INFORMATION: beta-1,4-galactosyltransferase (lgtB)
33 <400> SEQUENCE: 1
34 atg caa aac cac gtt atc agc tta gct tcc gcc gca gaa cgc agg gcg      48
35 Met Gln Asn His Val Ile Ser Leu Ala Ser Ala Ala Glu Arg Arg Ala
36   1               5                  10                  15
38 cac att gcc gat acc ttc ggc agg cac ggc atc ccg ttt cag ttt ttc      96
39 His Ile Ala Asp Thr Phe Gly Arg His Gly Ile Pro Phe Gln Phe Phe
40              20                  25                  30
42 gac gca ctg atg ccg tct gaa agg ctg gaa cag gca atg gcg gaa ctc     144
43 Asp Ala Leu Met Pro Ser Glu Arg Leu Glu Gln Ala Met Ala Glu Leu
44          35                  40                  45
46 gtc ccc ggc ttg tcg gcg cac ccc tat ttg agc gga gtg gaa aaa gcc     192
47 Val Pro Gly Leu Ser Ala His Pro Tyr Leu Ser Gly Val Glu Lys Ala
48      50                  55                  60
50 tgc ttt atg agc cac gcc gta ttg tgg aag cag gca ttg gac gaa ggt     240
51 Cys Phe Met Ser His Ala Val Leu Trp Lys Gln Ala Leu Asp Glu Gly
52  65                  70                  75                  80
54 ctg ccg tat atc acc gta ttt gag gac gac gtt tta ctc ggc gaa ggt     288
55 Leu Pro Tyr Ile Thr Val Phe Glu Asp Asp Val Leu Leu Gly Glu Gly
56              85                  90                  95
58 gag gaa aaa ttc ctt gcc gaa gac gct tgg ctg caa gaa cgc ttt gac     336
59 Glu Glu Lys Phe Leu Ala Glu Asp Ala Trp Leu Gln Glu Arg Phe Asp
60             100                 105                 110
62 ccg gat acc gcc ttt atc gtc cgc ttg gaa acg atg ttt atg cac gtc     384
63 Pro Asp Thr Ala Phe Ile Val Arg Leu Glu Thr Met Phe Met His Val
64             115                 120                 125
66 ctg acc tcg ccc tcc ggc gtg gcg gat tac tgc ggg cgc gcc ttt ccg     432
```

RAW SEQUENCE LISTING   DATE: 12/29/2000
PATENT APPLICATION: US/09/211,691   TIME: 14:01:06

Input Set : A:\-129-1.app
Output Set: N:\CRF3\12292000\I211691.raw

```
67 Leu Thr Ser Pro Ser Gly Val Ala Asp Tyr Cys Gly Arg Ala Phe Pro
68     130               135               140
70 ctg ttg gaa agc gaa cac tgg ggg acg gcg ggc tat atc att tcc cga    480
71 Leu Leu Glu Ser Glu His Trp Gly Thr Ala Gly Tyr Ile Ile Ser Arg
72 145               150               155               160
74 aaa gcg atg cgg ttt ttc ctg gac agg ttt gcc gcc ctg ccg ccc gaa    528
75 Lys Ala Met Arg Phe Phe Leu Asp Arg Phe Ala Ala Leu Pro Pro Glu
76             165               170               175
78 ggg ctg cac ccc gtc gat ctg atg atg ttc agc gat ttt ttc gac agg    576
79 Gly Leu His Pro Val Asp Leu Met Met Phe Ser Asp Phe Phe Asp Arg
80             180               185               190
82 gaa gga atg ccg gtt tgc cag ctc aat ccc gcc ttg tgc gcc caa gag    624
83 Glu Gly Met Pro Val Cys Gln Leu Asn Pro Ala Leu Cys Ala Gln Glu
84             195               200               205
86 ctg cat tat gcc aag ttt cac gac caa aac agc gca ttg ggc agc ctg    672
87 Leu His Tyr Ala Lys Phe His Asp Gln Asn Ser Ala Leu Gly Ser Leu
88     210               215               220
90 atc gaa cac gac cgc ctc ctg aac cgc aaa cag caa agg cgc gat tcc    720
91 Ile Glu His Asp Arg Leu Leu Asn Arg Lys Gln Gln Arg Arg Asp Ser
92 225               230               235               240
94 ccc gcc aac aca ttc aaa cac cgc ctg atc cgc gcc ttg acc aaa atc    768
95 Pro Ala Asn Thr Phe Lys His Arg Leu Ile Arg Ala Leu Thr Lys Ile
96             245               250               255
98 agc agg gaa agg gaa aaa cgc cgg caa agg cgc gaa cag ttc att gtg    816
99 Ser Arg Glu Arg Glu Lys Arg Arg Gln Arg Arg Glu Gln Phe Ile Val
100             260               265               270
102 cct ttc caa taa                                                    828
103 Pro Phe Gln
104     275
107 <210> SEQ ID NO: 2
108 <211> LENGTH: 275
109 <212> TYPE: PRT
110 <213> ORGANISM: Neisseria meningitidis
112 <400> SEQUENCE: 2
113 Met Gln Asn His Val Ile Ser Leu Ala Ser Ala Ala Glu Arg Arg Ala
114     1               5               10               15
115 His Ile Ala Asp Thr Phe Gly Arg His Gly Ile Pro Phe Gln Phe Phe
116             20               25               30
117 Asp Ala Leu Met Pro Ser Glu Arg Leu Glu Gln Ala Met Ala Glu Leu
118             35               40               45
119 Val Pro Gly Leu Ser Ala His Pro Tyr Leu Ser Gly Val Glu Lys Ala
120     50               55               60
121 Cys Phe Met Ser His Ala Val Leu Trp Lys Gln Ala Leu Asp Glu Gly
122 65               70               75               80
123 Leu Pro Tyr Ile Thr Val Phe Glu Asp Asp Val Leu Leu Gly Glu Gly
124             85               90               95
125 Glu Glu Lys Phe Leu Ala Glu Asp Ala Trp Leu Gln Glu Arg Phe Asp
126             100               105               110
127 Pro Asp Thr Ala Phe Ile Val Arg Leu Glu Thr Met Phe Met His Val
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/211,691

DATE: 12/29/2000
TIME: 14:01:06

Input Set : A:\-129-1.app
Output Set: N:\CRF3\12292000\I211691.raw

```
128          115                        120                     125
129 Leu Thr Ser Pro Ser Gly Val Ala Asp Tyr Cys Gly Arg Ala Phe Pro
130      130                      135              140
131 Leu Leu Glu Ser Glu His Trp Gly Thr Ala Gly Tyr Ile Ile Ser Arg
132 145                      150              155                  160
133 Lys Ala Met Arg Phe Phe Leu Asp Arg Phe Ala Ala Leu Pro Pro Glu
134              165                  170              175
135 Gly Leu His Pro Val Asp Leu Met Met Phe Ser Asp Phe Phe Asp Arg
136          180              185              190
137 Glu Gly Met Pro Val Cys Gln Leu Asn Pro Ala Leu Cys Ala Gln Glu
138          195              200              205
139 Leu His Tyr Ala Lys Phe His Asp Gln Asn Ser Ala Leu Gly Ser Leu
140      210              215              220
141 Ile Glu His Asp Arg Leu Leu Asn Arg Lys Gln Gln Arg Arg Asp Ser
142 225                  230              235                      240
143 Pro Ala Asn Thr Phe Lys His Arg Leu Ile Arg Ala Leu Thr Lys Ile
144              245              250              255
145 Ser Arg Glu Arg Glu Lys Arg Arg Gln Arg Arg Glu Gln Phe Ile Val
146          260              265              270
147 Pro Phe Gln
148          275
151 <210> SEQ ID NO: 3
152 <211> LENGTH: 41
153 <212> TYPE: DNA
154 <213> ORGANISM: Artificial Sequence
156 <220> FEATURE:
157 <223> OTHER INFORMATION: Description of Artificial Sequence:SYNTM-F1 5'
158      primer
160 <400> SEQUENCE: 3
161 cttaggaggt catatggaaa aacaaaatat tgcggttata c                41
164 <210> SEQ ID NO: 4
165 <211> LENGTH: 45
166 <212> TYPE: DNA
167 <213> ORGANISM: Artificial Sequence
169 <220> FEATURE:
170 <223> OTHER INFORMATION: Description of Artificial Sequence:SYNTM-R6 3'
171      primer
173 <400> SEQUENCE: 4
174 cgacagaatt ccgccaccgc tttccttgtg attaagaatg ttttc           45
177 <210> SEQ ID NO: 5
178 <211> LENGTH: 37
179 <212> TYPE: DNA
180 <213> ORGANISM: Artificial Sequence
182 <220> FEATURE:
183 <223> OTHER INFORMATION: Description of Artificial Sequence:SIALM-22F 5'
184      primer
186 <400> SEQUENCE: 5
187 gcatggaatt ctgggcttga aaaaggcttg tttgacc                    37
190 <210> SEQ ID NO: 6
```

```
191 <211> LENGTH: 59
192 <212> TYPE: DNA
193 <213> ORGANISM: Artificial Sequence
195 <220> FEATURE:
196 <223> OTHER INFORMATION: Description of Artificial Sequence:STALM-23R 3'
197      primer
199 <400> SEQUENCE: 6
200 cctaggtcga ctcattagtg gtgatggtgg tgatggttca ggtcttcttc gctgatcag  59
203 <210> SEQ ID NO: 7
204 <211> LENGTH: 9
205 <212> TYPE: PRT
206 <213> ORGANISM: Artificial Sequence
208 <220> FEATURE:
209 <223> OTHER INFORMATION: Description of Artificial Sequence:linker of
210      pFUS-01/2
212 <400> SEQUENCE: 7
213 Gly Gly Gly Ile Leu Ser His Gly Ile
214   1               5
217 <210> SEQ ID NO: 8
218 <211> LENGTH: 8
219 <212> TYPE: PRT
220 <213> ORGANISM: Artificial Sequence
222 <220> FEATURE:
223 <223> OTHER INFORMATION: Description of Artificial Sequence:linker of
224      pFUS-01/4
226 <400> SEQUENCE: 8
227 Gly Gly Gly Ile Leu Ser Gly Ile
228   1               5
231 <210> SEQ ID NO: 9
232 <211> LENGTH: 58
233 <212> TYPE: DNA
234 <213> ORGANISM: Artificial Sequence
236 <220> FEATURE:
237 <223> OTHER INFORMATION: Description of Artificial Sequence:GalE-5p 5'
238      primer
240 <400> SEQUENCE: 9
241 gggacaggat ccatcgatgc ttaggaggtc atatggcaat tttagtatta ggtggagc   58
244 <210> SEQ ID NO: 10
245 <211> LENGTH: 42
246 <212> TYPE: DNA
247 <213> ORGANISM: Artificial Sequence
249 <220> FEATURE:
250 <223> OTHER INFORMATION: Description of Artificial Sequence:GalE-3p 3'
251      primer
253 <400> SEQUENCE: 10
254 gggggggcta gcgccgcctc ctcgatcatc gtaccctttt gg                   42
257 <210> SEQ ID NO: 11
258 <211> LENGTH: 38
259 <212> TYPE: DNA
```

```
RAW SEQUENCE LISTING                    DATE: 12/29/2000
PATENT APPLICATION:   US/09/211,691     TIME: 14:01:06

Input Set : A:\-129-1.app
Output Set: N:\CRF3\12292000\I211691.raw
```

```
260 <213> ORGANISM: Artificial Sequence
262 <220> FEATURE:
263 <223> OTHER INFORMATION: Description of Artificial Sequence:LgtB-NheI 5'
264       primer
266 <400> SEQUENCE: 11
267 gggggggcta gcgtgcaaaa ccacgttatc agcttagc                        38
270 <210> SEQ ID NO: 12
271 <211> LENGTH: 45
272 <212> TYPE: DNA
273 <213> ORGANISM: Artificial Sequence
275 <220> FEATURE:
276 <223> OTHER INFORMATION: Description of Artificial Sequence:LgtB-SalI 3'
277       primer
279 <400> SEQUENCE: 12
280 gggggggtcg acctattatt ggaaaggcac aatgaactgt tcgcg                 45
283 <210> SEQ ID NO: 13
284 <211> LENGTH: 10
285 <212> TYPE: PRT
286 <213> ORGANISM: Artificial Sequence
288 <220> FEATURE:
289 <223> OTHER INFORMATION: Description of Artificial Sequence:peptide linker
291 <400> SEQUENCE: 13
292 Gly Gly Gly Ile Leu Ser His Gly Ile Leu
293   1               5                   10
296 <210> SEQ ID NO: 14
297 <211> LENGTH: 6
298 <212> TYPE: PRT
299 <213> ORGANISM: Artificial Sequence
301 <220> FEATURE:
302 <223> OTHER INFORMATION: Description of Artificial Sequence:6-His tail for
303       purification
305 <400> SEQUENCE: 14
306 His His His His His His
307   1               5
310 <210> SEQ ID NO: 15
311 <211> LENGTH: 5
312 <212> TYPE: PRT
313 <213> ORGANISM: Artificial Sequence
315 <220> FEATURE:
316 <223> OTHER INFORMATION: Description of Artificial Sequence:peptide linker
318 <400> SEQUENCE: 15
319 Gly Gly Ala Ser Val
320   1               5
323 <210> SEQ ID NO: 16
324 <211> LENGTH: 63
325 <212> TYPE: DNA
326 <213> ORGANISM: Artificial Sequence
328 <220> FEATURE:
329 <223> OTHER INFORMATION: Description of Artificial Sequence:junction region
```

VERIFICATION SUMMARY   DATE: 12/29/2000
PATENT APPLICATION: US/09/211,691 TIME: 14:01:07

Input Set : A:\-129-1.app
Output Set: N:\CRF3\12292000\I211691.raw

L:351 M:258 W: Mandatory Feature missing, <220> FEATURE: